

16643575 Oldon



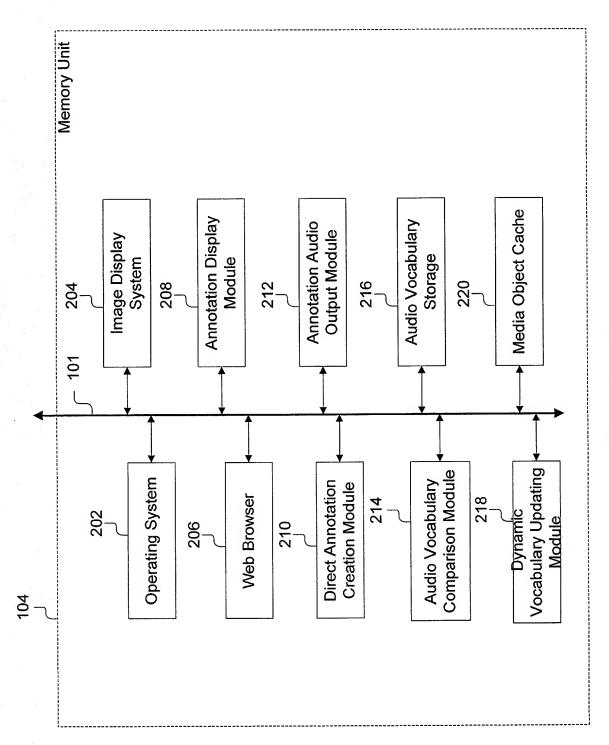
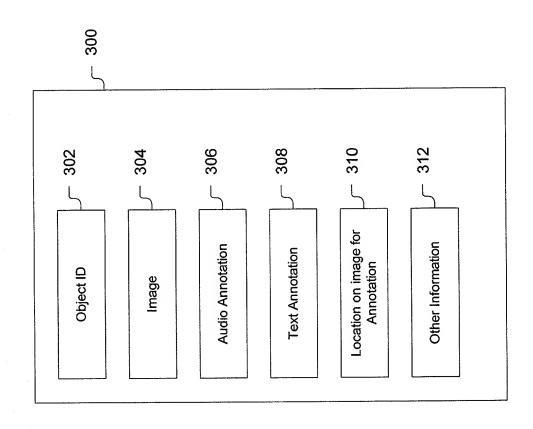


Figure 2



410
DISPLAY IMAGES

412
DETECT USER SELECTION OF AN IMAGE

414
RECORD AUDIO INPUT FROM USER

416
CREATE ANNOTATION BETWEEN
RECORDED AUDIO AND SELECTION

END

System and Method for Direct Multi-Modal

Sheet 4 of 20

Inventors: Gregory J. Wolff et al. Docket No. 20412-06364 Sh

## FIGURE 4

**START** 510 **DISPLAY IMAGES** 514 RECORD AUDIO INPUT FROM USER 512 DETECT USER SELECTION OF AN IMAGE 516 CREATE ANNOTATION BETWEEN RECORDED AUDIO AND SELECTION 518 DISPLAY IMAGES WITH SELECTED IMAGE HAVING AN VISUAL NOTATION OF AN **AUDIO ANNOTATION END** 

System and Method for Direct Multi-Modal

Sheet 5 of 20

Inventors: Gregory J. Wolff et al.

Docket No. 20412-06364

## FIGURE 5

**START** 610 **DISPLAY IMAGES** 612 DETECT USER SELECTION OF AN IMAGE 614 RECEIVE AUDIO INPUT FROM USER 616 COMPARE AUDIO INPUT TO A **VOCABULARY TO GENERATE TEXT** 618 **CREATE ANNOTATION BETWEEN** GENERATED TEXT AND SELECTION **END** 

System and Method for Direct Multi-Modal

Sheet 6 of 20

Inventors: Gregory J. Wolff et al.

Docket No. 20412-06364

## FIGURE 6

System and Method for Direct Multi-Modal Inventors: Gregory J. Wolff et al.

Docket No. 20412-06364 Sheet 7 of 20

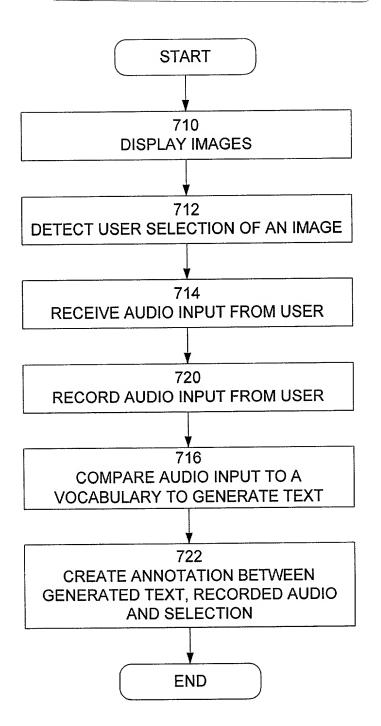


FIGURE 7

Docket No. 20412-06364 Sheet 8 of 20 **START** 810 **DISPLAY IMAGES** 812 DETECT USER SELECTION OF AN IMAGE 814 RECEIVE AUDIO INPUT FROM USER 816 **COMPARE AUDIO INPUT TO VOCABULARY** 818 **DOES AUDIO INPUT** MATCH? YES 820 CREATE ANNOTATION BETWEEN TEXT FOR MATCH AND SELECTION

System and Method for Direct Multi-Modal Inventors: Gregory J. Wolff et al.

## FIGURE 8A

**END** 

The last that the last the state of

Sheet 9 of 20

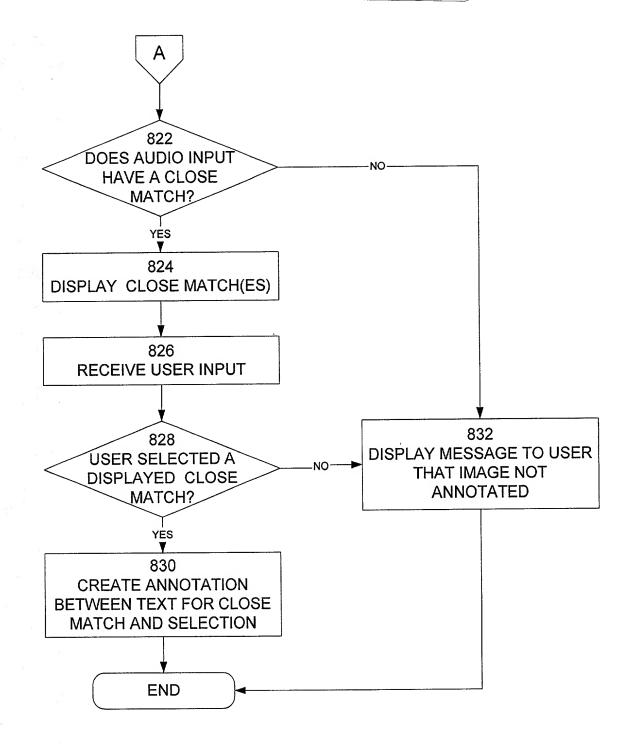


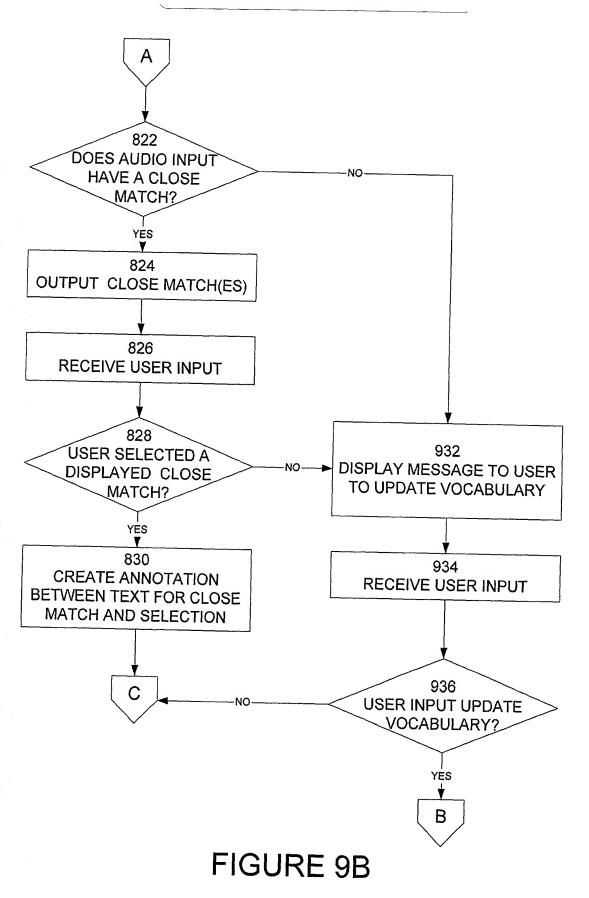
FIGURE 8B

System and Method for Direct Multi-Modal Inventors: Gregory J. Wolff et al. Docket No. 20412-06364 Sheet 10 of 20 **START** 810 **DISPLAY IMAGES** 812 DETECT USER SELECTION OF AN IMAGE 814 RECEIVE AUDIO INPUT FROM USER 816 **COMPARE AUDIO INPUT TO VOCABULARY** 818 DOES AUDIO INPUT MATCH? YĖS 820

FIGURE 9A

CREATE ANNOTATION BETWEEN TEXT FOR MATCH AND SELECTION





Docket No. 20412-06364 Sheet 12 of 20 В 938 REQUEST ANNOTATION FROM **USER** 940 RECEIVE USER INPUT 942 STORE TEXT AND AUDIO INPUT IN **VOCABULARY** 944 **CREATE ANNOTATION BETWEEN TEXT AND SELECTION** С 946 **DISPLAY IMAGES WITH SELECTED** IMAGE HAVING AN VISUAL NOTATION OF AN AUDIO **ANNOTATION** 

System and Method for Direct Multi-Modal

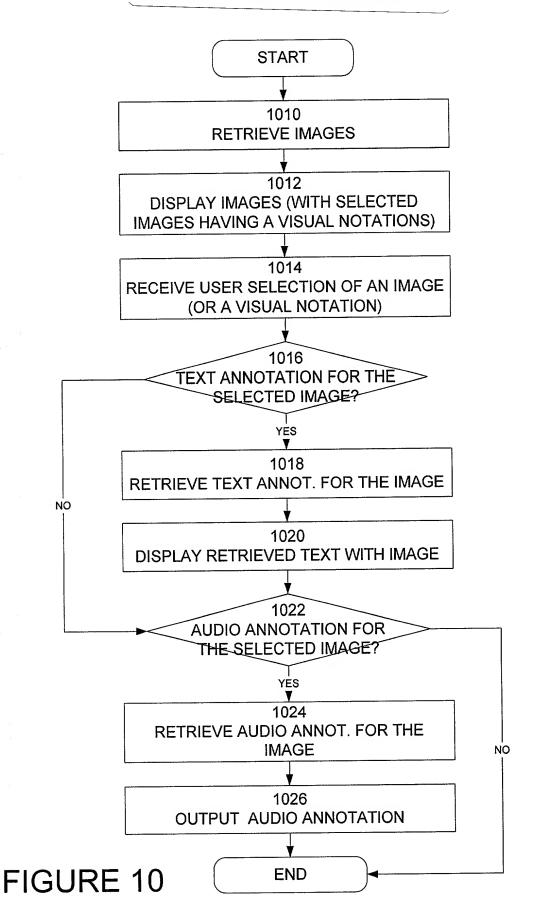
Inventors: Gregory J. Wolff et al.

FIGURE 9C

**END** 

System and Method for Direct Multi-Modal Inventors: Gregory J. Wolff et al.

Docket No. 20412-06364 Sheet 13 of 20



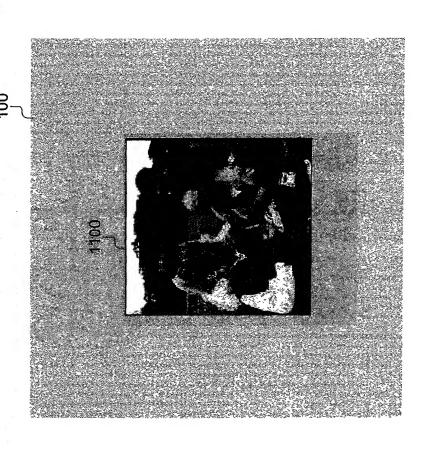
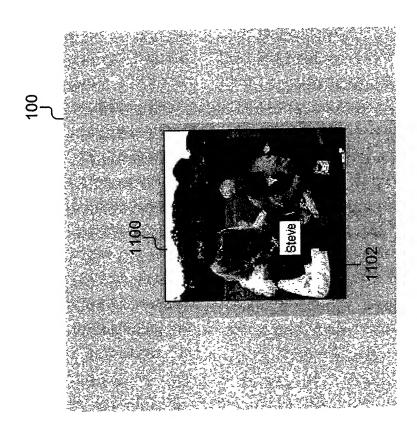


Figure 11A



IOOFZE/S.aliosce

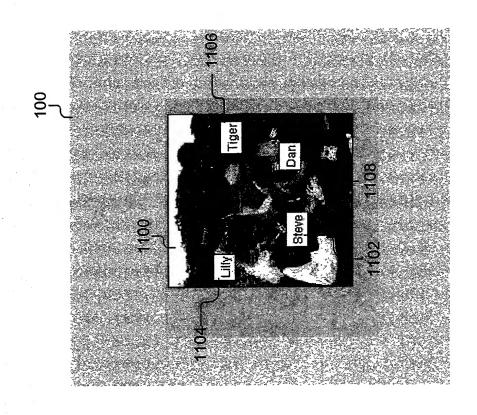


Figure 11C

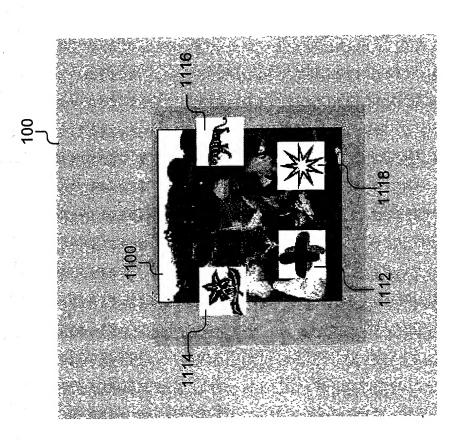


Figure 11D

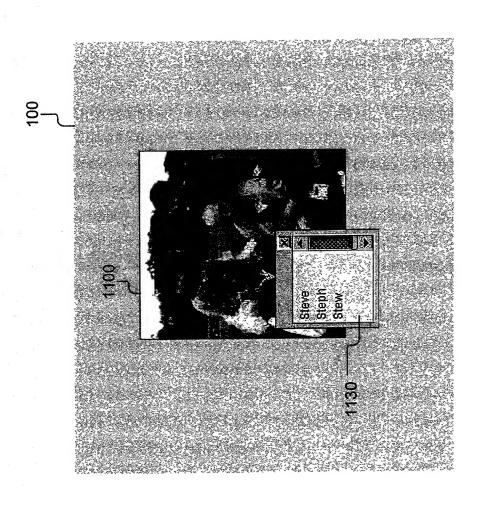
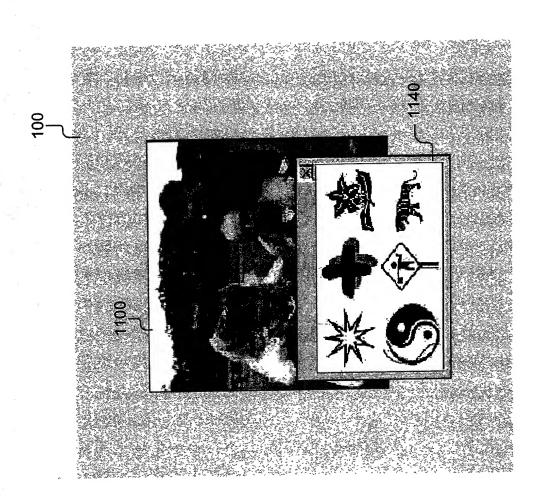


Figure 11E



Inventors: Gregory J. Wolff et al. Docket No. 20412-06364 Sheet 20 of 20 **START** 1210 **ENTER RETRIEVAL MODE** 1212 RECEIVE AUDIO INPUT FROM USER 1214 COMPARE AUDIO INPUT TO STORED **AUDIO SIGNALS** 1216 **DETERMINE CLOSEST MATCH(ES)** 1218 RETRIEVE IMAGE(S) REFERENCED BY **CLOSE MATCH ANNOTATION OBJECTS** 1220 DISPLAY RETRIEVED IMAGES

System and Method for Direct Multi-Modal

FIGURE 12

**END**